

O I P E  
MAY 19 1998

PATENT & TRADEMARK OFFICE  
SEARCHED INDEXED  
SERIALIZED FILED

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Luo, Yuling  
Xiomei, Xu

(ii) TITLE OF INVENTION: Semaphorin K1

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
(B) STREET: 75 DENISE DRIVE  
(C) CITY: HILLSBOROUGH  
(D) STATE: CALIFORNIA  
(E) COUNTRY: USA  
(F) ZIP: 94010

(v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/041,236  
(B) FILING DATE: March 11, 1998  
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: OSMAN, RICHARD A  
(B) REGISTRATION NUMBER: 36,627  
(C) REFERENCE/DOCKET NUMBER: EXEL98-001

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (650) 343-4341  
(B) TELEFAX: (650) 343-4342

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2498 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 1..1902

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTG	CTG	CTG	CTG	CTC	TGG	GCG	GCC	GCC	TCC	GCC	CAG	GGC	CAC	CTA	Leu	Leu	Leu	Leu	Leu	Trp	Ala	Ala	Ala	Ser	Ala	Gln	Gly	His	Leu	48		
1					5					10																	15					
AGG	AGC	GGA	CCC	CGC	ATC	TTC	GCC	GTC	TGG	AAA	GGC	CAT	GTA	GGG	CAG	Arg	Ser	Gly	Pro	Arg	Ile	Phe	Ala	Val	Trp	Lys	Gly	His	Val	96		
20																												30				
GAC	CGG	GTC	GAC	TTT	GGC	CAG	ACT	GAG	CCG	CAC	ACG	GTG	CTT	TTC	CAC	Asp	Arg	Val	Asp	Phe	Gly	Gln	Thr	Glu	Pro	His	Thr	Val	Leu	Phe	His	144
35																												45				
GAG	CCA	GGC	AGC	TCC	TCT	GTG	TGG	GGA	GGA	CGT	GGC	AAG	GTC	TAC	Glu	Pro	Gly	Ser	Ser	Val	Trp	Val	Gly	Arg	Gly	Lys	Val	Tyr	192			
50																												60				
CTC	TTT	GAC	TTC	CCC	GAG	GGC	AAG	AAC	GCA	TCT	GTG	CGC	ACG	GTG	AAT	Leu	Phe	Asp	Phe	Pro	Glu	Gly	Lys	Asn	Ala	Ser	Val	Arg	Thr	Val	Asn	240
65																												80				
ATC	GGC	TCC	ACA	AAG	GGG	TCC	TGT	CTG	GAT	AAG	CGG	GAC	TGC	GAG	AAC	Ile	Gly	Ser	Thr	Lys	Gly	Ser	Cys	Leu	Asp	Lys	Arg	Asp	Cys	Glu	Asn	288
85																												95				
TAC	ATC	ACT	CTC	CTG	GAG	AGG	CGG	AGT	GAG	GGG	CTG	CTG	GCC	TGT	GGC	Tyr	Ile	Thr	Leu	Leu	Glu	Arg	Arg	Ser	Glu	Gly	Leu	Leu	Ala	Cys	Gly	336
100																												110				
ACC	AAC	GCC	CGG	CAC	CCC	AGC	TGC	TGG	AAC	CTG	GTG	AAT	GGC	ACT	GTG	Thr	Asn	Ala	Arg	His	Pro	Ser	Cys	Trp	Asn	Leu	Val	Asn	Gly	Thr	Val	384

	115	120	125													
GTG	CCA	CTT	GCC	GAG	ATG	AGA	GGC	TAC	GCC	CCC	TTC	AGC	CCG	GAC	GAG	432
Val	Pro	Leu	Gly	Glu	Met	Arg	Gly	Tyr	Ala	Pro	Phe	Ser	Pro	Asp	Glu	
130					135						140					
AAC	TCC	CTG	GTT	CTG	TTT	GAA	GGG	GAC	GAG	GTG	TAT	TCC	ACC	ATC	CGG	480
Asn	Ser	Leu	Val	Leu	Phe	Glu	Gly	Asp	Glu	Val	Tyr	Ser	Thr	Ile	Arg	
145					150						155				160	
AAG	CAG	GAA	TAC	AAT	GGG	AAG	ATC	CCT	CGG	TTC	CGC	CGC	ATC	CGG	GGC	528
Lys	Gln	Glu	Tyr	Asn	Gly	Lys	Ile	Pro	Arg	Phe	Arg	Arg	Ile	Arg	Gly	
165					170						170				175	
GAG	AGT	GAG	CTG	TAC	ACC	AGT	GAT	ACT	GTC	ATG	CAG	AAC	CCA	CAG	TTC	576
Glu	Ser	Glu	Leu	Tyr	Thr	Ser	Asp	Thr	Val	Met	Gln	Asn	Pro	Gln	Phe	
180					185						185				190	
ATC	AAA	GCC	ACC	ATC	GTG	CAC	CAA	GAC	CAG	GCT	TAC	GAT	GAC	AAG	ATC	624
Ile	Lys	Ala	Thr	Ile	Val	His	Gln	Asp	Gln	Ala	Tyr	Asp	Asp	Lys	Ile	
195					200						200				205	
TAC	TAC	TTC	TTC	CGA	GAG	GAC	AAT	CCT	GAC	AAG	AAT	CCT	GAG	GCT	CCT	672
Tyr	Tyr	Phe	Phe	Arg	Glu	Asp	Asn	Pro	Asp	Lys	Asn	Pro	Glu	Ala	Pro	
210					215						220					
CTC	AAT	GTG	TCC	CGT	GTG	GCC	CAG	TTG	TGC	AGG	GGG	GAC	CAG	GGT	GGG	720
Leu	Asn	Val	Ser	Arg	Val	Ala	Gln	Leu	Cys	Arg	Gly	Asp	Gln	Gly	Gly	
225					230						235				240	
GAA	AGT	TCA	CTG	TCA	GTC	TCC	AAG	TGG	AAC	ACT	TTT	CTG	AAA	GCC	ATG	768
Glu	Ser	Ser	Leu	Ser	Ser	Val	Ser	Lys	Trp	Asn	Thr	Phe	Leu	Lys	Ala	Met
245					250						250				255	
CTG	GTA	TGC	AGT	GAT	GCT	GCC	ACC	AAC	AAG	AAC	TTC	AAC	AGG	CTG	CAA	816
Leu	Val	Cys	Ser	Asp	Ala	Ala	Thr	Asn	Lys	Asn	Phe	Asn	Arg	Leu	Gln	
260					265						270					
GAC	GTC	TTC	CTG	CTC	CCT	GAC	CCC	AGC	GGC	CAG	TGG	AGG	GAC	ACC	AGG	864
Asp	Val	Phe	Leu	Leu	Pro	Asp	Pro	Ser	Gly	Gln	Trp	Arg	Asp	Thr	Arg	
275					280						285					
GTC	TAT	GGT	GGT	TTC	TTC	AAC	CCC	TGG	AAC	TAC	TCA	GCC	GTC	TGT	GTG	912
Val	Tyr	Gly	Val	Phe	Ser	Asn	Pro	Trp	Asn	Tyr	Ser	Ala	Val	Cys	Val	
290					295						300					
TAT	TCC	CTC	GGT	GAC	ATT	GAC	AAG	GTC	TTC	CGT	ACC	TCC	TCA	CTC	AAG	960
Tyr	Ser	Leu	Gly	Asp	Ile	Asp	Lys	Val	Phe	Arg	Thr	Ser	Ser	Leu	Lys	
305					310						315				320	
GGC	TAC	CAC	TCA	AGC	CTT	CCC	AAC	CCG	CGG	CCT	GGC	AAG	TGC	CTC	CCA	1008
Gly	Tyr	His	Ser	Ser	Leu	Pro	Asn	Pro	Arg	Pro	Gly	Lys	Cys	Leu	Pro	
325					330						330				335	
GAC	CAG	CAG	CCG	ATA	CCC	ACA	GAG	ACC	TTC	CAG	GTG	GCT	GAC	CGT	CAC	1056
Asp	Gln	Gln	Pro	Ile	Pro	Thr	Glu	Thr	Phe	Gln	Val	Ala	Asp	Arg	His	
340					345						350					
CCA	GAG	GTG	GCG	CAG	AGG	GTG	GAG	CCC	ATG	GGG	CCT	CTG	AAG	ACG	CCA	1104
Pro	Glu	Val	Ala	Gln	Arg	Val	Glu	Pro	Met	Gly	Pro	Leu	Lys	Thr	Pro	
355					360						365					
TTG	TTC	CAC	TCT	AAA	TAC	CAC	TAC	CAG	AAA	GTG	GCC	GTC	CAC	CGC	ATG	1152
Leu	Phe	His	Ser	Lys	Tyr	His	Tyr	Gln	Lys	Val	Ala	Val	His	Arg	Met	
370					375						380					
CAA	GCC	AGC	CAC	GGG	GAG	ACC	TTT	CAT	GTG	CTT	TAC	CTA	ACT	ACA	GAC	1200
Gln	Ala	Ser	His	Gly	Glu	Thr	Phe	His	Val	Leu	Tyr	Leu	Thr	Thr	Asp	
385					390						395				400	
AGG	GGC	ACT	ATC	CAC	AAG	GTG	GTG	GAA	CCG	GGG	GAG	CAG	GAG	CAC	AGC	1248
Arg	Gly	Thr	Ile	His	Lys	Val	Val	Glu	Pro	Gly	Glu	Gln	Glu	His	Ser	
405					410						410				415	
TTC	GCC	TTC	AAC	ATC	ATG	GAG	ATC	CAG	CCC	TTC	CGC	CGC	GCG	GCT	GCC	1296
Phe	Ala	Phe	Asn	Ile	Met	Glu	Ile	Gln	Pro	Phe	Arg	Arg	Ala	Ala	Ala	
420					425						430					
ATC	CAG	ACC	ATG	TCG	CTG	GAT	GCT	GAG	CGG	AGG	AAG	CTG	TAT	GTG	AGC	1344
Ile	Gln	Thr	Met	Ser	Leu	Asp	Ala	Glu	Arg	Arg	Lys	Leu	Tyr	Val	Ser	
435					440						445					
TCC	CAG	TGG	GAG	GTG	AGC	CAG	GTG	CCC	CTG	GAC	CTG	TGT	GAG	GTC	TAT	1392
Ser	Gln	Trp	Glu	Val	Ser	Gln	Val	Pro	Leu	Asp	Leu	Cys	Glu	Val	Tyr	
450					455						460					

GGC GGG GGC TGC CAC GGT TGC CTC ATG TCC CGA GAC CCC TAC TGC GGC	1440
Gly Gly Cys His Gly Cys Leu Met Ser Arg Asp Pro Tyr Cys Gly	
465 470 475 480	
TGG GAC CAA GGC CGC TGC ATC TCC ATC TAC AGC TCC GAA CGG TCA GTG	1488
Trp Asp Gln Gly Arg Cys Ile Ser Ile Tyr Ser Ser Glu Arg Ser Val	
485 490 495	
CTG CAA TCC ATT AAT CCA GCC GAG CCA CAC AAG GAG TGT CCC AAC CCC	1536
Leu Gln Ser Ile Asn Pro Ala Glu Pro His Lys Glu Cys Pro Asn Pro	
500 505 510	
AAA CCA GAC AAG GCC CCA CTG CAG AAG GTT TCC CTG GCC CCA AAC TCT	1584
Lys Pro Asp Lys Ala Pro Leu Gln Lys Val Ser Leu Ala Pro Asn Ser	
515 520 525	
CGC TAC TAC CTG AGC TGC CCC ATG GAA TCC CGC CAC GCC ACC TAC TCA	1632
Arg Tyr Tyr Leu Ser Cys Pro Met Glu Ser Arg His Ala Thr Tyr Ser	
530 535 540	
TGG CGC CAC AAG GAG AAC GTG GAG CAG AGC TGC GAA CCT GGT CAC CAG	1680
Trp Arg His Lys Glu Asn Val Glu Gln Ser Cys Glu Pro Gly His Gln	
545 550 555 560	
AGC CCC AAC TGC ATC CTG TTC ATC GAG AAC CTC ACG GCG CAG CAG TAC	1728
Ser Pro Asn Cys Ile Leu Phe Ile Glu Asn Leu Thr Ala Gln Gln Tyr	
565 570 575	
GGC CAC TAC TTC TGC GAG GCC CAG GAG GGC TCC TAC TTC CGC GAG GCT	1776
Gly His Tyr Phe Cys Glu Ala Gln Glu Gly Ser Tyr Phe Arg Glu Ala	
580 585 590	
CAG CAC TGG CAG CTG CTG CCC GAG GAC GGC ATC ATG GCC GAG CAC CTG	1824
Gln His Trp Gln Leu Leu Pro Glu Asp Gly Ile Met Ala Glu His Leu	
595 600 605	
CTG GGT CAT GCC TGT GCC CTG GCC TCC CTC TGG CTG GGG GTG CTG	1872
Leu Gly His Ala Cys Ala Leu Ala Ala Ser Leu Trp Leu Gly Val Leu	
610 615 620	
CCC ACA CTC ACT CTT GGC TTG CTG GTC CAC TAGGGCCTCC CGAGGCTGGG	1922
Pro Thr Leu Thr Leu Gly Leu Leu Val His	
625 630	
CATGCCCTCAG GCTTCTGCAG CCCAGGGCAC TAAAACGTCT CACACTCAGA GCCGGCTGGC	1982
CCGGGAGCTC CTTGCTGCC ATTTTTTCCA GGGGACAGAA TAACCCAGTG GAGGATGCCA	2042
GGCCTGGAGA CGTCCAGCCG CAGGGCGCTG CTGGGGCCCA GGTGGCGCAC GGATGGTGAG	2102
GGGCTGAGAA TGAGGGCACC GACTGTGAAG CTGGGGCATC GATGACCCAA GACTTTATT	2162
TTTGGAAAAT ATTTTCAGA CTCCCTCAAAC TTGACTAAAT GCAGCGATGC TCCCCAGCCCA	2222
AGAGCCCAGT GGTCGGGGAG TGGGTTTGGA TAGGAGAGCT GGGATTCCAT CTCGACCCCTG	2282
GGGCTGAGGC CTGAGTCCTT TTGGATTCTT GGTACCCACA TTGCCTCCTT CCCCTCCTTT	2342
TTTCAGGGGT GGGTGGTTGG TGTTCCTGAA GACCCAGGGGA TACCCTTTGT CCAGCCCTGT	2402
CCTTGGCAGC TCCCTTTTG GTCCCTGGTC CCACAGGACA GCCGCCTTGC ATGTTTATTG	2462
AAGGATGTTT GCTTCCCGGA CGGAAGGACG GAAAAA	2498

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 634 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Leu Leu Leu Leu Trp Ala Ala Ala Ser Ala Gln Gly His Leu	
1 5 10 15	
Arg Ser Gly Pro Arg Ile Phe Ala Val Trp Lys Gly His Val Gly Gln	
20 25 30	
Asp Arg Val Asp Phe Gly Gln Thr Glu Pro His Thr Val Leu Phe His	
35 40 45	
Glu Pro Gly Ser Ser Ser Val Trp Val Gly Gly Arg Gly Lys Val Tyr	
50 55 60	
Leu Phe Asp Phe Pro Glu Gly Lys Asn Ala Ser Val Arg Thr Val Asn	
65 70 75 80	
Ile Gly Ser Thr Lys Gly Ser Cys Leu Asp Lys Arg Asp Cys Glu Asn	
85 90 95	
Tyr Ile Thr Leu Leu Glu Arg Arg Ser Glu Gly Leu Leu Ala Cys Gly	

100                   105                   110  
Thr Asn Ala Arg His Pro Ser Cys Trp Asn Leu Val Asn Gly Thr Val  
115                   120                   125  
Val Pro Leu Gly Glu Met Arg Gly Tyr Ala Pro Phe Ser Pro Asp Glu  
130                   135                   140  
Asn Ser Leu Val Leu Phe Glu Gly Asp Glu Val Tyr Ser Thr Ile Arg  
145                   150                   155                   160  
Lys Gln Glu Tyr Asn Gly Lys Ile Pro Arg Phe Arg Arg Ile Arg Gly  
165                   170                   175  
Glu Ser Glu Leu Tyr Thr Ser Asp Thr Val Met Gln Asn Pro Gln Phe  
180                   185                   190  
Ile Lys Ala Thr Ile Val His Gln Asp Gln Ala Tyr Asp Asp Lys Ile  
195                   200                   205  
Tyr Tyr Phe Phe Arg Glu Asp Asn Pro Asp Lys Asn Pro Glu Ala Pro  
210                   215                   220  
Leu Asn Val Ser Arg Val Ala Gln Leu Cys Arg Gly Asp Gln Gly Gly  
225                   230                   235                   240  
Glu Ser Ser Leu Ser Val Ser Lys Trp Asn Thr Phe Leu Lys Ala Met  
245                   250                   255  
Leu Val Cys Ser Asp Ala Ala Thr Asn Lys Asn Phe Asn Arg Leu Gln  
260                   265                   270  
Asp Val Phe Leu Leu Pro Asp Pro Ser Gly Gln Trp Arg Asp Thr Arg  
275                   280                   285  
Val Tyr Gly Val Phe Ser Asn Pro Trp Asn Tyr Ser Ala Val Cys Val  
290                   295                   300  
Tyr Ser Leu Gly Asp Ile Asp Lys Val Phe Arg Thr Ser Ser Leu Lys  
305                   310                   315                   320  
Gly Tyr His Ser Ser Leu Pro Asn Pro Arg Pro Gly Lys Cys Leu Pro  
325                   330                   335  
Asp Gln Gln Pro Ile Pro Thr Glu Thr Phe Gln Val Ala Asp Arg His  
340                   345                   350  
Pro Glu Val Ala Gln Arg Val Glu Pro Met Gly Pro Leu Lys Thr Pro  
355                   360                   365  
Leu Phe His Ser Lys Tyr His Tyr Gln Lys Val Ala Val His Arg Met  
370                   375                   380  
Gln Ala Ser His Gly Glu Thr Phe His Val Leu Tyr Leu Thr Thr Asp  
385                   390                   395                   400  
Arg Gly Thr Ile His Lys Val Val Glu Pro Gly Glu Gln Glu His Ser  
405                   410                   415  
Phe Ala Phe Asn Ile Met Glu Ile Gln Pro Phe Arg Arg Ala Ala Ala  
420                   425                   430  
Ile Gln Thr Met Ser Leu Asp Ala Glu Arg Arg Lys Leu Tyr Val Ser  
435                   440                   445  
Ser Gln Trp Glu Val Ser Gln Val Pro Leu Asp Leu Cys Glu Val Tyr  
450                   455                   460  
Gly Gly Gly Cys His Gly Cys Leu Met Ser Arg Asp Pro Tyr Cys Gly  
465                   470                   475                   480  
Trp Asp Gln Gly Arg Cys Ile Ser Ile Tyr Ser Ser Glu Arg Ser Val  
485                   490                   495  
Leu Gln Ser Ile Asn Pro Ala Glu Pro His Lys Glu Cys Pro Asn Pro  
500                   505                   510  
Lys Pro Asp Lys Ala Pro Leu Gln Lys Val Ser Leu Ala Pro Asn Ser  
515                   520                   525  
Arg Tyr Tyr Leu Ser Cys Pro Met Glu Ser Arg His Ala Thr Tyr Ser  
530                   535                   540  
Trp Arg His Lys Glu Asn Val Glu Gln Ser Cys Glu Pro Gly His Gln  
545                   550                   555                   560  
Ser Pro Asn Cys Ile Leu Phe Ile Glu Asn Leu Thr Ala Gln Gln Tyr  
565                   570                   575  
Gly His Tyr Phe Cys Glu Ala Gln Glu Gly Ser Tyr Phe Arg Glu Ala  
580                   585                   590  
Gln His Trp Gln Leu Leu Pro Glu Asp Gly Ile Met Ala Glu His Leu  
595                   600                   605  
Leu Gly His Ala Cys Ala Leu Ala Ala Ser Leu Trp Leu Gly Val Leu

610                    615                    620  
 Pro Thr Leu Thr Leu Gly Leu Leu Val His  
 625                    630

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1818 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..1818

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG	GGC	ACT	TTG	TGT	GTT	AGT	ATT	AGA	TTA	CTG	ATG	ATT	TTA	TCA	GCC	48
Met	Gly	Thr	Leu	Cys	Val	Ser	Ile	Arg	Leu	Leu	Met	Ile	Leu	Ser	Ala	
635				640					645				650			
ATC	ACA	GCT	GCT	AAA	TCT	CGG	TTT	ATA	GAT	AAG	CCA	AGG	CTG	ATT	GTT	96
Ile	Thr	Ala	Ala	Lys	Ser	Arg	Phe	Ile	Asp	Lys	Pro	Arg	Leu	Ile	Val	
				655				660				665				
AAC	CTA	ACT	GAT	GGG	TTT	GGA	CAG	CAC	CGG	TTT	TTT	GGA	CCC	CAG	GAA	144
Asn	Leu	Thr	Asp	Gly	Phe	Gly	Gln	His	Arg	Phe	Phe	Gly	Pro	Gln	Glu	
				670			675				680					
CCA	CAC	ACT	GTG	CTT	TTT	CAC	AGC	CTC	AAC	TCT	TCA	GAC	GTA	TAT	GTG	192
Pro	His	Thr	Val	Leu	Phe	His	Ser	Leu	Asn	Ser	Ser	Asp	Val	Tyr	Val	
				685			690				695					
GGA	GGT	AAT	AAT	ACC	ATC	TAT	TTG	TTT	GAT	TTT	GCT	CAC	AGC	TCC	AAC	240
Gly	Gly	Asn	Asn	Thr	Ile	Tyr	Leu	Phe	Asp	Phe	Ala	His	Ser	Ser	Asn	
				700			705				710					
GCA	TCC	ACA	GCT	TTG	ATA	AAC	ATA	ACT	AGC	ACA	CAT	AAT	ACC	CAC	CGG	288
Ala	Ser	Thr	Ala	Leu	Ile	Asn	Ile	Thr	Ser	Thr	His	Asn	Thr	His	Arg	
				715			720				725			730		
TTA	TCT	AGT	ACC	TGC	GAA	AAC	TTT	ATA	ACT	CTG	CTT	CAT	AAC	CAG	ACA	336
Leu	Ser	Ser	Thr	Cys	Glu	Asn	Phe	Ile	Thr	Leu	Leu	His	Asn	Gln	Thr	
				735			740				745					
GAT	GGG	CTG	CTA	GCT	TGT	GGT	ACT	AAC	TCA	CAG	AAA	CCC	AGC	TGC	TGG	384
Asp	Gly	Leu	Leu	Ala	Cys	Gly	Thr	Asn	Ser	Gln	Lys	Pro	Ser	Cys	Trp	
				750			755				760					
CTG	ATA	AAC	AAC	CTA	ACA	ACT	CAA	TTT	TTG	GGG	CCA	AAA	CTA	GGC	TTA	432
Leu	Ile	Asn	Asn	Leu	Thr	Thr	Gln	Phe	Leu	Gly	Pro	Lys	Leu	Gly	Leu	
				765			770				775					
GCC	CCC	TTC	TCA	CCA	TCA	TCT	GGC	AAT	CTG	GTG	CTG	TTT	GAC	CAG	AAT	480
Ala	Pro	Phe	Ser	Pro	Ser	Ser	Gly	Asn	Leu	Val	Leu	Phe	Asp	Gln	Asn	
				780			785				790					
GAC	ACC	TAT	TCC	ACC	ATT	AAC	CTC	TAC	AAG	AGC	CTC	AGT	GGC	TCT	CAC	528
Asp	Thr	Tyr	Ser	Thr	Ile	Asn	Leu	Tyr	Lys	Ser	Leu	Ser	Gly	Ser	His	
				795			800				805			810		
AAG	TTT	AGG	AGG	ATC	GCT	GGC	CAA	GTA	GAA	CTA	TAC	ACG	AGT	GAC	ACC	576
Lys	Phe	Arg	Arg	Ile	Ala	Gly	Gln	Val	Glu	Leu	Tyr	Thr	Ser	Asp	Thr	
				815			820				825					
GCC	ATG	CAC	CGG	CCA	CAG	TTT	GTC	CAG	GCA	ACA	GCT	GTG	CAT	AAA	AAT	624
Ala	Met	His	Arg	Pro	Gln	Phe	Val	Gln	Ala	Thr	Ala	Val	His	Lys	Asn	
				830			835				840					
GAA	TCT	TAT	GAT	GAT	AAA	ATC	TAC	TTT	TTC	TTT	CAA	GAA	AAC	AGC	CAC	672
Glu	Ser	Tyr	Asp	Asp	Lys	Ile	Tyr	Phe	Phe	Phe	Gln	Glu	Asn	Ser	His	
				845			850				855					
AGT	GAC	TTC	AAA	CAG	TTT	CCA	CAT	ACT	GTA	CCT	AGA	GTG	GGG	CAG	GTG	720
Ser	Asp	Phe	Lys	Gln	Phe	Pro	His	Thr	Val	Pro	Arg	Val	Gly	Gln	Val	
				860			865				870					
TGC	TCT	AGT	GAT	CAA	GGT	GGG	GAG	AGC	TCC	CTG	TCT	GTC	TAC	AAG	TGG	768
Cys	Ser	Ser	Asp	Gln	Gly	Gly	Glu	Ser	Ser	Leu	Ser	Val	Tyr	Lys	Trp	
				875			880				885			890		
ACC	ACC	TTT	TTA	AAA	GCC	AGA	CTG	GCG	TGT	GTA	GAC	TAT	GAT	ACT	GGA	816

Thr	Thr	Phe	Leu	Lys	Ala	Arg	Leu	Ala	Cys	Val	Asp	Tyr	Asp	Thr	Gly	
		895			900				905							
AGA	ATC	TAC	AAT	GAG	CTA	CAA	GAT	ATT	TTC	ATC	TGG	CAA	GCC	CCA	GAG	864
Arg	Ile	Tyr	Asn	Glu	Leu	Gln	Asp	Ile	Phe	Ile	Trp	Gln	Ala	Pro	Glu	
		910			915				920							
AAC	AGC	TGG	GAA	GAG	ACT	CTC	ATC	TAT	GGA	CTT	TTT	TTG	AGC	CCG	TGG	912
Asn	Ser	Trp	Glu	Glu	Thr	Leu	Ile	Tyr	Gly	Leu	Phe	Leu	Ser	Pro	Trp	
		925			930				935							
AAC	TTT	TCT	GCG	GTC	TGT	G TG	TTT	ACT	GTA	AAG	GAC	ATT	GAC	CAT	GTG	960
Asn	Phe	Ser	Ala	Val	Cys	Val	Phe	Thr	Val	Lys	Asp	Ile	Asp	His	Val	
		940			945				950							
TTT	AAG	ACA	TCC	AAG	TTA	AAA	AAT	TAT	CAT	CAT	AAA	CTC	CCC	ACA	CCT	1008
Phe	Lys	Thr	Ser	Lys	Leu	Lys	Asn	Tyr	His	His	Lys	Leu	Pro	Thr	Pro	
		955			960				965			970				
AGA	CCA	GGG	CAA	TGC	ATG	AAG	AAC	CAT	CAG	CAT	GTT	CCC	ACA	GAA	ACC	1056
Arg	Pro	Gly	Gln	Cys	Met	Lys	Asn	His	Gln	His	Val	Pro	Thr	Glu	Thr	
		975			980				985							
TTT	CAG	GTT	GCT	GAC	AGA	TAT	CCA	GAA	GTT	GCA	GAT	CCT	GTA	TAT	CAG	1104
Phe	Gln	Val	Ala	Asp	Arg	Tyr	Pro	Glu	Val	Ala	Asp	Pro	Val	Tyr	Gln	
		990			995				1000							
AAG	AAC	AAT	GCC	ATG	TTT	CCA	ATA	ATT	CAG	TCA	AAA	TAT	ATC	TAC	ACC	1152
Lys	Asn	Asn	Ala	Met	Phe	Pro	Ile	Ile	Gln	Ser	Lys	Tyr	Ile	Tyr	Thr	
		1005			1010				1015							
AAA	CTA	CTT	GTT	TAT	AGG	GTA	GAG	TAT	GGA	GGT	GTT	TTT	TGG	GCA	ACT	1200
Lys	Leu	Leu	Val	Tyr	Arg	Val	Glu	Tyr	Gly	Gly	Val	Phe	Trp	Ala	Thr	
		1020			1025				1030							
ATT	TTT	TAC	CTC	ACT	ACC	ATC	AAA	GGG	ACT	ATT	CAT	ATA	TAT	GTG	AGG	1248
Ile	Phe	Tyr	Leu	Thr	Thr	Ile	Lys	Gly	Thr	Ile	His	Ile	Tyr	Val	Arg	
		1035			1040				1045			1050				
TAT	GAA	GAT	TCC	AAC	TCT	ACA	ACA	GCT	CTC	AAC	ATT	TTA	GAG	ATA	AAT	1296
Tyr	Glu	Asp	Ser	Asn	Ser	Thr	Thr	Ala	Leu	Asn	Ile	Leu	Glu	Ile	Asn	
		1055			1060				1065							
CCC	TTT	CAG	AAG	CCA	GCC	CCC	ATA	CAG	AAT	ATT	CTT	TTA	GAT	AAT	ACA	1344
Pro	Phe	Gln	Pro	Ala	Pro	Ile	Gln	Asn	Ile	Leu	Leu	Asp	Asn	Thr		
		1070			1075				1080							
AAT	CTA	AAG	CTT	TAT	GTA	AAT	TCA	GAG	TGG	GAG	GTG	AGT	GAG	GTG	CCA	1392
Asn	Leu	Lys	Leu	Tyr	Val	Asn	Ser	Glu	Trp	Glu	Val	Ser	Glu	Val	Pro	
		1085			1090				1095							
TTA	GAC	CTA	TGT	TCA	GTG	TAT	GGG	AAT	GAT	TGT	TTC	AGC	TGT	TTT	ATG	1440
Leu	Asp	Leu	Cys	Ser	Val	Tyr	Gly	Asn	Asp	Cys	Phe	Ser	Cys	Phe	Met	
		1100			1105				1110							
TCA	AGG	GAT	CCC	CTG	TGC	ACA	TGG	TAT	AAC	AAC	ACC	TGT	TCC	TTT	AAA	1488
Ser	Arg	Asp	Pro	Leu	Cys	Thr	Trp	Tyr	Asn	Asn	Thr	Cys	Ser	Phe	Lys	
		1115			1120				1125			1130				
CAG	AGA	GTA	TCT	GTT	GAA	ACC	GGT	GGT	CCA	GCT	AAC	CGC	ACC	CTT	TCA	1536
Gln	Arg	Val	Ser	Val	Glu	Thr	Gly	Gly	Pro	Ala	Asn	Arg	Thr	Leu	Ser	
		1135			1140				1145							
GAA	ATG	TGT	GGT	GAC	CAC	TAT	GCT	CCA	ACT	GTG	GTT	AAG	CAT	CAA	GTT	1584
Glu	Met	Cys	Gly	Asp	His	Tyr	Ala	Pro	Thr	Val	Val	Lys	His	Gln	Val	
		1150			1155				1160							
TCT	ATA	CCT	CTA	TTA	TCT	AAT	TCT	TAT	TTG	TCC	TGC	CCA	GCA	GTC	TCA	1632
Ser	Ile	Pro	Leu	Leu	Ser	Asn	Ser	Tyr	Leu	Ser	Cys	Pro	Ala	Val	Ser	
		1165			1170				1175							
AAC	CAC	GCT	GAC	TAC	TTT	TGG	ACT	AAA	GAT	GGT	TTC	ACA	GAA	AAA	AGA	1680
Asn	His	Ala	Asp	Tyr	Phe	Trp	Thr	Lys	Asp	Gly	Phe	Thr	Glu	Lys	Arg	
		1180			1185				1190							
TGC	CAT	GTC	AAA	ACA	CAC	AAA	AAT	GAC	TGC	ATC	TTG	CTT	ATA	GCT	AAC	1728
Cys	His	Val	Lys	Thr	His	Lys	Asn	Asp	Cys	Ile	Leu	Leu	Ile	Ala	Asn	
		1195			1200				1205			1210				
AGC	ACG	ACA	GCC	ACT	AAT	GGA	ACC	CAC	GTG	TGC	AAC	ATG	AAA	GAA	GAT	1776
Ser	Thr	Thr	Ala	Thr	Asn	Gly	Thr	His	Val	Cys	Asn	Met	Lys	Glu	Asp	
		1215			1220				1225							
TCG	GTG	ACA	GTG	AAA	CTG	TTA	GAG	GTG	AAT	GTG	ACA	CTG	ATG			1818
Ser	Val	Thr	Val	Lys	Leu	Leu	Glu	Val	Asn	Val	Thr	Leu	Met			

1230

1235

1240

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 606 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gly Thr Leu Cys Val Ser Ile Arg Leu Leu Met Ile Leu Ser Ala  
 1 5 10 15

Ile Thr Ala Ala Lys Ser Arg Phe Ile Asp Lys Pro Arg Leu Ile Val  
 20 25 30

Asn Leu Thr Asp Gly Phe Gly Gln His Arg Phe Phe Gly Pro Gln Glu  
 35 40 45

Pro His Thr Val Leu Phe His Ser Leu Asn Ser Ser Asp Val Tyr Val  
 50 55 60

Gly Gly Asn Asn Thr Ile Tyr Leu Phe Asp Phe Ala His Ser Ser Asn  
 65 70 75 80

Ala Ser Thr Ala Leu Ile Asn Ile Thr Ser Thr His Asn Thr His Arg  
 85 90 95

Leu Ser Ser Thr Cys Glu Asn Phe Ile Thr Leu Leu His Asn Gln Thr  
 100 105 110

Asp Gly Leu Leu Ala Cys Gly Thr Asn Ser Gln Lys Pro Ser Cys Trp  
 115 120 125

Leu Ile Asn Asn Leu Thr Thr Gln Phe Leu Gly Pro Lys Leu Gly Leu  
 130 135 140

Ala Pro Phe Ser Pro Ser Ser Gly Asn Leu Val Leu Phe Asp Gln Asn  
 145 150 155 160

Asp Thr Tyr Ser Thr Ile Asn Leu Tyr Lys Ser Leu Ser Gly Ser His  
 165 170 175

Lys Phe Arg Arg Ile Ala Gly Gln Val Glu Leu Tyr Thr Ser Asp Thr  
 180 185 190

Ala Met His Arg Pro Gln Phe Val Gln Ala Thr Ala Val His Lys Asn  
 195 200 205

Glu Ser Tyr Asp Asp Lys Ile Tyr Phe Phe Gln Glu Asn Ser His  
 210 215 220

Ser Asp Phe Lys Gln Phe Pro His Thr Val Pro Arg Val Gly Gln Val  
 225 230 235 240

Cys Ser Ser Asp Gln Gly Gly Glu Ser Ser Leu Ser Val Tyr Lys Trp  
 245 250 255

Thr Thr Phe Leu Lys Ala Arg Leu Ala Cys Val Asp Tyr Asp Thr Gly  
 260 265 270

Arg Ile Tyr Asn Glu Leu Gln Asp Ile Phe Ile Trp Gln Ala Pro Glu  
 275 280 285

Asn Ser Trp Glu Glu Thr Leu Ile Tyr Gly Leu Phe Leu Ser Pro Trp  
 290 295 300

Asn Phe Ser Ala Val Cys Val Phe Thr Val Lys Asp Ile Asp His Val  
 305 310 315 320

Phe Lys Thr Ser Lys Leu Lys Asn Tyr His His Lys Leu Pro Thr Pro  
 325 330 335

Arg Pro Gly Gln Cys Met Lys Asn His Gln His Val Pro Thr Glu Thr  
 340 345 350

Phe Gln Val Ala Asp Arg Tyr Pro Glu Val Ala Asp Pro Val Tyr Gln  
 355 360 365

Lys Asn Asn Ala Met Phe Pro Ile Ile Gln Ser Lys Tyr Ile Tyr Thr  
 370 375 380

Lys Leu Leu Val Tyr Arg Val Glu Tyr Gly Gly Val Phe Trp Ala Thr  
 385 390 395 400

Ile Phe Tyr Leu Thr Thr Ile Lys Gly Thr Ile His Ile Tyr Val Arg  
 405 410 415

Tyr Glu Asp Ser Asn Ser Thr Thr Ala Leu Asn Ile Leu Glu Ile Asn  
 420 425 430

Pro Phe Gln Lys Pro Ala Pro Ile Gln Asn Ile Leu Leu Asp Asn Thr

	435	440	445												
Asn	Leu	Lys	Leu	Tyr	Val	Asn	Ser	Glu	Trp	Glu	Val	Ser	Glu	Val	Pro
	450			455						460					
Leu	Asp	Leu	Cys	Ser	Val	Tyr	Gly	Asn	Asp	Cys	Phe	Ser	Cys	Phe	Met
465				470						475					480
Ser	Arg	Asp	Pro	Leu	Cys	Thr	Trp	Tyr	Asn	Asn	Thr	Cys	Ser	Phe	Lys
				485					490						495
Gln	Arg	Val	Ser	Val	Glu	Thr	Gly	Gly	Pro	Ala	Asn	Arg	Thr	Leu	Ser
				500				505							510
Glu	Met	Cys	Gly	Asp	His	Tyr	Ala	Pro	Thr	Val	Val	Lys	His	Gln	Val
	515				520							525			
Ser	Ile	Pro	Leu	Leu	Ser	Asn	Ser	Tyr	Leu	Ser	Cys	Pro	Ala	Val	Ser
	530				535						540				
Asn	His	Ala	Asp	Tyr	Phe	Trp	Thr	Lys	Asp	Gly	Phe	Thr	Glu	Lys	Arg
545					550					555					560
Cys	His	Val	Lys	Thr	His	Lys	Asn	Asp	Cys	Ile	Leu	Leu	Ile	Ala	Asn
					565				570						575
Ser	Thr	Thr	Ala	Thr	Asn	Gly	Thr	His	Val	Cys	Asn	Met	Lys	Glu	Asp
					580				585						590
Ser	Val	Thr	Val	Lys	Leu	Leu	Glu	Val	Asn	Val	Thr	Leu	Met		
					595			600							605